Amendments to the Specification:

Please replace paragraph [76] with the following amended paragraph:

[76] Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al., Nuc. Acids Res.* **25**:3389-3402 (1977) and Altschul *et al., J. Mol. Biol.* **215**:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (~~**http://www.ncbi.nlm.nih.gov/**~~ **URL address: http file type, www host server, domain name ncbi.nlm.gov**). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al., supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA*

89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

Please replace paragraph [166] with the following amended paragraph:

[166] In other preferred embodiments, an assay such as the fluorescence polarization assay or the fluorescence resonance energy transfer assay is employed to identify candidate modulators. These assays do not require the separation of bound and free labeled test compound. Fluorescence polarization (FP) or fluorescence anisotropy is a useful tool for the study of molecular interactions (*see, e.g.* ~~**http://www.panvera.com/tech/ appguide/fpintro.html**~~ **URL address: http file type, www host server, domain name panvera.com/tech/ appguide/fpintro.html**, November 4, 1999). First, a molecule labeled with a fluorophore is excited with plane polarized light. If the fluorescent molecule stays stationary while in the excited state, light is emitted in the same polarized plane. If the excited fluorescently labeled molecule rotates out of the plane of the polarized light while in the excited state, light is emitted from the molecule in a different plane. For example, if vertical polarized light is used to excite the fluorophore, the emission spectra can be monitored in the vertical and horizontal planes. Fluorescence polarization is calculated as shown in the following Formula I:

Please replace paragraph [233] with the following amended paragraph:

[233] A full-length cDNA corresponding to AA237916 was isolated from a mouse liver cDNA library (Origene), and this sequence was used to identify a human ortholog in the GENBANK EST database (T86384). A full-length human sequence was obtained by iterative EST database searches and by cloning from human liver cDNA libraries (Origene and Clontech). The human cDNA predicts a 651 amino acid protein (FIGURE 1C) that shares 82% sequence identity with the mouse sequence (data not shown). Following the standard system of nomenclature in the ABC transporter field, this protein has been named *ABCG5*. The amino terminal half of *ABCG5* contains the ATP-binding motifs (Walker A and B motifs) and an ABC

transporter signature motif (C motif), while the carboxyl terminal region is predicted to contain six transmembrane (TM) segments (FIGURE 1B) (C. Higgins, *Annu. Rev. Cell Biol.* **8**:67 (1992) Walker *et al., Embo. J.* **1**:945 (1982) Jones *et al., Biochem.* **33**:3038 (1994)). A human EST clone (Unigene T93792) from *ABCG5* mapped to chromosome 2p21 between markers D2S177 and D2S119 and the map position was confirmed using a radiation hybrid panel. Chromosomal localization of *ABCG5* was confirmed using primers derived from exon 7 of *ABCG5* to amplify a gene-specific fragment from the TNG panel of radiation hybrids from Stanford Human Genome Center (Research Genetics, Inc). The result was submitted to the RH Server (~~**http://wwwshgc.stanford.edu/cgi-bin**~~ **URL address: http file type, www host server, domain name -shgc.stanford.edu**), which linked *ABCG5* to SHGC14952, which is between markers D2S177 and D2S119. Patel and colleagues previously mapped the human sitosterolemia gene to this same region of chromosome 2 in ten independent sitosterolemic families (Patel *et al., J. Clin. Invest.* **102**:1041 (1998)).

Please replace paragraph [235] with the following amended paragraph:

[235] Genes encoding members of the ABC transporter family are often clustered in the genome (Le Saux *et al., Nat. Genet.* **25**:223 (2000)). Since only a single *ABCG5* mutation was identified in our collection of nine sitosterolemic probands, the public and Celera genome sequences were searched for other ABC transporters adjacent to *ABCG5*. An EST (T84531) that shared weak homology with the *Drosophila white* gene was identified and expanded using exons predicted by the computer program GENSCAN. The 3'-end of *ABCG5* was located on BAC RP11-489K22, which had been partially sequenced, but no other ABC transporters were identified on this BAC. A BAC end sequences (BES) in the Genome Survey Sequence database that was located on BAC RP11-489K22 was used to search the Celera Human Fragments database. The public and Celera database were used to assemble most of the genomic sequences in the region, resulting in the identification of EST T84531, which shared weak homology with the *drosophila white* gene (Repa *et al., Science* **289**:1524 (2000)). The

GENSCAN Web Server (~~http://genes.mit.edu/GENSCAN.html~~ **URL address: http file type, www host server, domain name genes.mit.edu/GENSCAN.html**) was used to identify additional exons within this gene. The sequence of the ~30 kb region was assembled (excluding three gaps) using the Celera Human Fragments database and mouse ESTs in the public database. Eleven of the 13 exons of the new gene, which was named *ABCG8*, were identified in the databases and the remaining two exons were identified by sequencing PCR-amplified cDNAs from human hepatic polyA$^+$ mRNA. *ABCG8* shares ~28% amino acid identity with *ABCG5* (FIGURE 1C). Its sequence is most similar to *ABCG1*, which resembles the *Drosophila white* gene (Bingham *et al., Cell,* **25**:693 (1981)).